A Case Study of IV&V Return on Investment (ROI)

Dr. Richard A. Rogers , AverStar Inc., Greenbelt MD
Dan McCaugherty, AverStar Inc., Fairmont WV
Dr. Fred Martin, AverStar Inc., Burlington MA

Abstract

This study determines the return on investment (ROI) for IV&V based on a comparison of two groups of computer software configuration items (CSCIs) associated with Space Shuttle ground systems. One group received IV&V services across the full development cycle, the other beginning in the code and development test phase. Observed differences in error densities between the two groups indicate that those CSCIs receiving full life cycle IV&V had fewer defects during code and development test. Industry data on the costs associated with finding and fixing these defects, combined with actual IV&V costs, are used to determine an effective ROI of between 1.25 and 1.82 for the IV&V. For reasons elucidated in the paper these results, good as they are, actually understate the full value of IV&V.

Background

The CSCIs comprising the data set for the study came from two Space Shuttle ground systems projects at Johnson Space Center in the early 90's: the Day of Launch I-Load Update (DOLILU) project and the Flight Analysis and Design System (FADS) project. DOLILU software determined, validated and up-linked Space Shuttle first stage guidance commands based on environmental and atmospheric data collected in the hours prior to launch. FADS took the DOLILU software and redesigned and rebuilt it for distributed UNIX workstations.

Different contractors, using different development regimes, developed individual DOLILU and FADS CSCIs. AverStar Inc., at that time operating as Intermetrics, Inc., was contracted to perform Independent Verification and Validation (IV&V) on both projects. IV&V analyses distinguished five phases of development: requirements, architecture design, detailed design, code development and test, and formal test. IV&V was applied starting at different points in the development cycle. In some cases, IV&V was not initiated until software designs were complete and in other cases requirements and design artifacts were so limited as to preclude IV&V for those phases. Records were kept of the defects found by IV&V in each lifecycle phase and the IV&V level of effort applied to each CSCI.

The study authors were able to categorize CSCIs based on the extent of the IV&V services applied and the timing of those services. The "Full Lifecycle IV&V" group is comprised of CSCIs for which IV&V analyses were performed and defects tabulated during all five phases of the development cycle. The CSCIs in this group are as follows:

For DOLILU – DIBS, DTS, RSOLNK SVDS and PROC
For FADS – TLAMS, DTS and DIVPLT

The "No Requirements and/or Design IV&V" group is comprised of CSCIs that either did not have IV&V performed or did not have the artifacts necessary to enable IV&V during at least one of first three development phases. These CSCI did not have the benefit of having defects identified for one or more of requirements, architecture design or detailed design. The net effect being that defects normally identified by IV&V in these phases were still present during code and development test and subsequent phases. CSCIs in this group are as follows:

For DOLILU – DIVDT, ILDMP, ILUV and LTQS
For FADS – Scripts

Defects for each group of CSCIs were tabulated with the results summarized in the table below. The data show the Full Life Cycle IV&V group had 133.3 fewer defects per 1000 function points during the code and development test phase.

Group Identification	Number of Function Points	Total Number of Defects During Development & Test	Development & Test Defects per 1K Function Points
Full Lifecycle IV&V	3482	237	68 per 1K FP
No Rqmts/Design IV&V	1832	369	201 per 1K FP

These data provide the basis for a determination of the ROI associated with IV&V defined as:

$$(1) \qquad \text{ROI} = \frac{\text{Value of IV\&V}}{\text{Cost of IV\&V}}$$

Both Value and Cost can be measured either in terms of the level of effort (LOE) or dollars expended. LOE measured as person months was chosen because the age of the data sets brought into question the relevancy of costs measured in dollars and LOE information is more prevalent than dollars in the literature.

Cost of IV&V

A total of 49.75 person months of effort was expended performing IV&V for the CSCIs in the Full Lifecycle IV&V group across all five phases of the development effort. This equates to 14.3 person months of IV&V per 1K function points.

Breakouts of labor by individual lifecycle phase were not available. One study author who participated in the IV&V estimates the percentage expended on requirements and design IV&V (the first three phases) as 35% to 45%. Analysis of data taken from a Logicon chart included in an Air Force System Command publication on IV&V [AFSC] determines this range as 37% to 53%. Using the range of 40% to 50% for the proportion of total effort expended during requirements and design IV&V puts the cost of this IV&V in the range of 5.7 to 7.2 person months per 1K function point.

Value of IV&V

The value of IV&V is embodied in the 133.3 fewer defects per 1K function points in code that has received early lifecycle IV&V. This establishes the number of defects that would need to be identified and repaired during code and test had they not been detected and corrected earlier. The value of IV&V becomes the cost savings associated with not having to identify these defects plus the net savings in repair costs. The net savings in repair costs result from the fact that the earlier in the life cycle a problem is corrected the cheaper it is to do so. This can be expressed as

$$(2) \qquad \text{Value of IV\&V} = 133.3 * (\text{Defect ID Costs} + \text{Defect Repair Costs Differential}),$$

Where:

Defect ID Costs = the cost of identifying a software defect during the code and development test phase, and

Defect Repair Costs Differential = the differential in costs associated with repairing a defect during the code and development phase as compared with repairs made during requirements and design.

Defect identification (ID) costs during test were determined by averaging the data presented in Table 9.1, page 276 of Watts Humphrey's *A Discipline for Software Engineering* [Humphrey] to get 6.8 to 8.5 hours per defect.

The defect repair costs differential was computed as the difference between the repair rates for testing defect removal and non-test defect removal as reported in Table 1 page 175 of Capers Jones' book *Software Quality: Analysis and Guidelines for Success* [Jones]. Based on the data in the mode column of this table the defect repair costs differential was determined to be 4 hours per defect.

Converting to person months (160 hours per month) and applying equation (2) determines the value of IV&V to be in the range of 9 to 10.4 person months per 1K function points.

Conclusion and Summary Discussion

Applying equation (1) across the ranges established above for IV&V value and costs yields:

$$(3) \qquad 1.25 \leq \text{IV\&V ROI} \leq 1.82$$

The defect identification costs used in this computation include fixed costs associated with the test environment (equipment purchase, equipment maintenance, facility expenses, and fixed operating costs). In a strict sense these costs would be incurred regardless of whether IV&V had or had not been performed and should be excluded from the savings being attributed to IV&V. A 10% reduction in IV&V value and therefore in IV&V ROI would probably overstate the impact of fixed costs associated with testing.

There are other factors that would tend to indicate that the values in (3) represent an underestimate of the IV&V ROI. Chief among these is the fact that the above analysis has tacitly assumed that all 133.3 defects would be detected during code and development test. Industry data do not support this assumption. Data from the chapter on Defect Removal Efficiency [Jones] indicate that a 5% defect leakage rate (i.e., testing finds 95% of the defects) would be a conservative estimate. This means that of the 133.3 defects present during code and development test, at least 6.5 can be expected to survive into formal acceptance testing and beyond where the cost to repair can increase substantially. Sources in the literature [Boehm, Humphrey] report 4- to 10-fold increases in defect repair costs during later phases of the life cycle as compared to earlier phases.

The value of IV&V also is increased by the "watchdog effect" (having IV&V on the job makes the developer more diligent) and by cumulative effects during maintenance associated with having better, more complete and more accurate documentation for the software. Finally, the worth of IV&V for mission critical software often goes beyond purely economic considerations. In cases where a single undetected fault can disrupt or negate a mission the value of IV&V as a risk reduction "insurance policy" can transcend the economics on which this analysis is based.

References

[AFSC] Air Force Systems Command, *Software Independent Verification and Validation (IV&V)*, AFSC AFLC Pamphlet 800-5, 1988

[Boehm] Boehm, Barry w., *Software Engineering Economics*, Prentice-Hall, Englewood Cliffs, NJ, 1981

[Humphrey] Humphrey, Watts S., *A Discipline for Software Engineering*, Addison-Wesley Publishing Company, Reading, MA, 1995

[Jones] Jones, Capers, *Software Quality: Analysis and Guidelines for Success*, International Thomson Computer Press, Boston, MA, 1997